RNA-seq analysis of bronchial organoids reveals ACE2 expression levels before and after SARS-CoV-2 infection

Edward Y. Chen

Master of Science in Biology

Division of biological Sciences, University of California San Diego, La Jolla, CA, USA

ABSTRACT

The SARS-CoV-2 virus invades human cells via the angiotensin-converting enzyme 2 (ACE2) receptor. ACE2 expression in the lung cells creates a target for the SARS-CoV-2 virus, and consequently infected individuals often see respiratory symptoms. Here, RNA sequencing (RNA-seq) datasets from *in vitro* human bronchial organoids (hBO) are analyzed to identify underlying gene expression changes as a result of an infection by SARS-CoV-2. Through differential expression analysis, it was found that ACE2 expression levels did not change significantly after SARS-CoV-2 infection, but other genes that encode for major immune signaling pathways were upregulated. This study can lend insight into potential targets for drug development and treatment for the COVID-19 disease.

INTRODUCTION

The emergence of the novel coronavirus (SARS-CoV-2) in late 2019 has caused the COVID-19 global pandemic. This virus comes from the family *Coronaviridae*, and previous forms of the coronavirus have caused the SARS (SARS-CoV) and the MERS (MERS-CoV) epidemics.¹⁰ This

family of RNA virus is known to cause pneumonia-like respiratory symptoms such as coughing and shortness of breath.²¹ The SARS-CoV-2 (as well as the previous SARS-CoV) virus contains 3-D spike proteins and is known to have a high binding affinity to ACE2.²⁰

ACE2 is found in several types of organ tissues, and its role in the renin-angiotensin system (RAS) is to cleave angiotensin 2 (AngII) to angiotensin 1-7.^{6,24} AngII is a peptide hormone that causes vasoconstriction and higher blood pressure.⁶ ACE2 essentially down-regulates RAS by inactivating AngII.

This tight correlation of the coronavirus with the ACE2 receptor is substantiated by a study that showed a positive correlation between the level of ACE2 expression and SARS-CoV-2 infection susceptibility.⁷ ACE2 expression levels in the cell are often determined by the stage of cell differentiation, and the more differentiated a cell is, the more it expresses ACE2.¹²

In addition, it is found that the receptor ACE2 is necessary for viral entry into the cell, and the virus's receptor binding domain (RBD) specifically binds to the Lysine 31 residue of ACE2.²² The SARS-CoV-2 virus enters the host cell through binding and endocytosis with ACE2 (instead of membrane fusion), effectively removing ACE2 from cell surfaces.^{9,19}

This study will use bioinformatics to determine the effects of SARS-CoV-2 on the gene expression levels of ACE2 using publicly available RNA sequencing data. It is predicted that this removal from the cell surface after viral binding will decrease ACE2 levels. This decrease in ACE2 will lead to an increase in AngII, and such imbalance in the RAS is known show symptoms consistent with COVID-19 such as heart failure, fibrosis and vasoconstriction.⁹ This inflammatory response can further create a positive feedback mechanism due to the involvement

of the transmembrane protease ADAM-17 which is responsible for activating cytokines and shedding ACE2.^{2,3}

The RNA-seq samples are from an *in vitro* generation of bronchial organoids from commercially available human bronchial epithelial cells. Differential expression analysis can provide insight into how the expressions of genes are altered to due an infection. Also, this analysis can reveal if gene expression of ACE2 decreases after SARS-CoV-2 infection. RNA-seq analysis can also aid in the development of drugs or medicines to target and block certain receptors to mitigate the symptoms of this disease.

MATERIALS AND METHODS

The bulk RNA sequencing data were obtained from the public NCBI website under the GEO database (<https://trace.ncbi.nlm.nih.gov/Traces/sra/?study=SRP262285>). The samples were human bronchial organoids generated from cryopreserved and commercially available human adult airway epithelial cells.¹⁸ Three samples were uninfected (SRR11811019, SRR11811020, SRR11811021), and three samples (SRR11811022, SRR11811023, SRR11811024) were infected with the SARS-CoV-2 virus *in vitro*.

The bulk RNA sequencing data were single-ended and already trimmed before downloading into the terminal. After downloading, the SRR files were ran through the FastQC software (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) for quality control check.² The reference transcriptome data for *homo sapiens* was obtained from the kallisto software homepage (<https://github.com/pachterlab/kallisto-transcriptome-indices/releases>). Then, the individual RNA sequencing reads were mapped to the reference transcriptome using the kallisto

pseudo-alignment software.⁵ After the alignment, the gene counts from the abundance files were used to generate a matrix in Microsoft Excel 2018.

For differential expression analysis, the DESeq2 software¹⁶ for the R programming language (<http://www.bioconductor.org/packages/release/bioc/html/DESeq2.html>) was used to normalize the data, and a log transformation was applied as a correction for multiple testing. The quantified gene counts were used to determine the statistical significance for each gene between the two treatments. Using the list of significant genes ($p < 0.01$) generated by DESeq2, the software Metascape²³ was used (<http://metascape.org>) to identify which of the significant genes were the most highly expressed after infection.

RESULTS

All reads were of good quality after the FastQC check, and each sequence length was 101 base pairs long. After the kallisto mapping step, the uninfected samples were on average 88% pseudo-aligned. The infected samples were on average 89% pseudo-aligned. The output matrix contained 188753 different genes. Through differential analysis, it was found that 1953 genes were significantly different across treatments with a p-values less than 0.01.

The gene expression counts of ACE2 splice variant 1 (Fig.1) did not differ between uninfected and infected samples ($p=0.95$). The gene expression counts of ACE2 splice variant 2 (Fig.2) also did not differ between uninfected and infected treatments ($p=0.98$). The heatmap generated shows that the overall gene expression of the infected samples is upregulated compared to that of the uninfected samples (Fig. 3). This pattern is substantiated by the principal component analysis (Fig.4) whose variance revealed that the samples were inherently different when it came to the

expression of its gene due to the two different clusters that were generated. PC1 was able to explain 55% of the total variance between gene expressions of the samples. Although there is one outlier in one of the infected samples along PC2.

Figure 1. Normalized expression counts for ACE2 gene splice variant 1.

Figure 2. Normalized expression counts for ACE2 gene splice variant 2.

Figure 3. Heatmap of the gene expression profile before and after infection

Figure 4. Principal component analysis of the differentially expressed genes grouped by treatment.

The Metascape analysis was performed using the list of significantly different genes with $p < 0.01$ (Fig.4). The GO term assigned to the largest number of significantly expressed genes were those involved in cytokine mediated signaling pathways (GO:0019221). The second largest group of genes were those involved in regulated exocytosis (GO:0045055).

Figure 4. Gene ontology analysis of the functions of the significant genes ($p < 0.01$) and their corresponding p-values.

DISCUSSION

As a result of the infection by SARS-CoV-2, certain genes in the bronchial organoids have shown significant changes in their expressions. However, ACE2 expression remained the same contrary to the original hypothesis. 1953 out of the total 188753 genes were expressed differently ($p < 0.01$) as a result of the viral infection. The genes that were significantly different were the ones involved in the inflammatory and immune responses. The maintenance of ACE2 may play a protective role as ACE2 has been shown to mediate certain forms of lung injuries.^{11,14}

The differential expression was demonstrated with the principal component analysis which showed the cells in different treatments clustering together, showing their variation in gene expressions before and after infection. However, along PC2, there was one outlier in the infected samples. This could be due to the nature of stem cell experimentation. Organoids have the advantage of being cell cultures that can capture the 3-D architecture of the *in vivo* cells and their functions, but contaminants could have been introduced during the sampling process.

The two splice variants of ACE2 did not show significant differences before and after SARS-CoV-2 infection. The similar p-values between the splice variants could mean that the virus is able to couple with each variant.⁷ It remains to be studied if the virus can decouple from its receptor if more variants arise in ACE2.

That the ACE2 expression remained the same could mean that ACE2 might not respond directly to the viral entry, or it does not contribute to the immune or inflammatory response in these particular samples. One study by Li and colleagues¹⁵ showed that ACE2 positively correlated with CD8+ T cells in males infected with SARS-CoV-2, but is negatively correlated in females. This means that the association of ACE2 with immune signatures could vary across individuals. Other complexities involved could be the ages of the individuals in which the samples were taken, which for this study, is unknown. One study described the paradoxical accounts of the geriatric population who are more susceptible to SARS-CoV-2, but have less ACE2 receptors expressed compared to the pediatric population.²

Future studies could compare ACE2 expression between different types of lung tissue, or even comparing among organs to test their sensitivity to SARS-CoV-2. This could lend insight into

other symptoms experienced by COVID-19 patients. The *in vivo* lung is highly differentiated, and it would be hard to fully capture that complexity *in vitro*. More repetitions with more samples are definitely necessary to come to a more robust conclusion.

Based on the Metascape analysis, cytotoxicity and the immune response were observed as a result of SARS-CoV-2 infection. Cytokines are proteins that bind to surface receptors that trigger downstream cellular pathways such as transcription. Cytokines are usually released as a response to some stimuli. This can lead to the inflammatory response and help recruit immune cells and other signaling molecules such as interferons.⁴ SARS-CoV-2 activates toll-like receptors and pro-inflammatory cytokines including interleukins (IL)-1 β , IL-6 and IL-8 that mediate lung inflammation, fever, and fibrosis.⁸

Another future direction could be to investigate the role of RAS in response to SARS-CoV-2 as well as determine the gene expressions of AngII in response to SARS-CoV-2. Another future direction could be investigating the role of the renin-angiotensin system on the inflammatory response, and to see if an imbalance in other enzymes or proteins in this system could worsen lung damage or cause symptoms.

ACKNOWLEDGEMENTS

Thank you to Suzuki and colleagues who provided the online RNA sequencing data. Thank you to the reviewers of this paper.

REFERENCES

1. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Internet]. Babraham Bioinformatics - FastQC A Quality Control tool for High Throughput Sequence Data. 2010 [cited 2020Aug4]. Available from: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>
2. Behl T, Kaur I, Bungau S, Kumar A, Uddin MS, Kumar C, et al. The dual impact of ACE2 in COVID-19 and ironical actions in geriatrics and pediatrics with possible therapeutic solutions. *Life Sciences*. 2020;257:118075.
3. Black RA, Rauch CT, Kozlosky CJ, Peschon JJ, Slack JL, Wolfson MF, et al. A metalloproteinase disintegrin that releases tumour-necrosis factor- α from cells. *Nature*. 1997;385(6618):729–33.
4. Borthwick LA, Wynn TA, Fisher AJ. Cytokine mediated tissue fibrosis. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*. 2013;1832(7):1049–60.
5. Bray NL, Pimentel H, Melsted P, Pachter L. Erratum: Near-optimal probabilistic RNA-seq quantification. *Nature Biotechnology*. 2016;34(8):888–.
6. Burrell LM, Johnston CI, Tikellis C, Cooper ME. ACE2, a new regulator of the renin–angiotensin system. *Trends in Endocrinology & Metabolism*. 2004;15(4):166–9.
7. Cao Y, Li L, Feng Z, Wan S, Huang P, Sun X, et al. Comparative genetic analysis of the novel coronavirus (2019-nCoV/SARS-CoV-2) receptor ACE2 in different populations. *Cell Discovery*. 2020;6(1).
8. Conti P, Ronconi G, Caraffa A, Gallenga CE, Ross R, Frydas I, et al. Induction of pro-inflammatory cytokines (IL-1 and IL-6) and lung inflammation by Coronavirus-19 (COVID-19 or SARS-CoV-2):

Anti-inflammatory strategies. *Journal of Biological Regulators & Homeostatic agents*.

2020Mar26;34(2).

9. Gheblawi M, Wang K, Viveiros A, Nguyen Q, Zhong J-C, Turner AJ, et al. Angiotensin-Converting Enzyme 2: SARS-CoV-2 Receptor and Regulator of the Renin-Angiotensin System. *Circulation Research*. 2020;126(10):1456–74.
10. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell*. 2020;181(2).
11. Imai Y, Kuba K, Rao S, Huan Y, Guo F, Guan B, et al. Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature*. 2005;436(7047):112–6.
12. Jia HP, Look DC, Shi L, Hickey M, Pewe L, Netland J, et al. ACE2 Receptor Expression and Severe Acute Respiratory Syndrome Coronavirus Infection Depend on Differentiation of Human Airway Epithelia. *Journal of Virology*. 2005;79(23):14614–21.
13. Jia HP, Look DC, Tan P, Shi L, Hickey M, Gakhar L, et al. Ectodomain shedding of angiotensin converting enzyme 2 in human airway epithelia. *American Journal of Physiology-Lung Cellular and Molecular Physiology*. 2009;297(1).
14. Kuba K, Imai Y, Rao S, Gao H, Guo F, Guan B, et al. A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus-induced lung injury. *Nature Medicine*. 2005;11(8):875–9.
15. Li M-Y, Li L, Zhang Y, Wang X-S. Expression of the SARS-CoV-2 cell receptor gene ACE2 in a wide variety of human tissues. *Infectious Diseases of Poverty*. 2020;9(1).

16. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*. 2014;15(12).
17. Shenoy V, Kwon K-C, Rathinasabapathy A, Lin S, Jin G, Song C, et al. Oral Delivery of Angiotensin-Converting Enzyme 2 and Angiotensin-(1-7) Bioencapsulated in Plant Cells Attenuates Pulmonary Hypertension. *Hypertension*. 2014;64(6):1248–59.
18. Suzuki T, Itoh Y, Sakai Y, Saito A, Okuzaki D, Motooka D, et al. Generation of human bronchial organoids for SARS-CoV-2 research. 2020;
19. Wang H, Yang P, Liu K, Guo F, Zhang Y, Zhang G, et al. SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. *Cell Research*. 2008;18(2):290–301.
20. Xu H, Zhong L, Deng J, Peng J, Dan H, Zeng X, et al. High expression of ACE2 receptor of 2019-nCoV on the epithelial cells of oral mucosa. *International Journal of Oral Science*. 2020;12(1).
21. Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*. 2020;367(6485):1444–8.
22. Zhang H, Penninger JM, Li Y, Zhong N, Slutsky AS. Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. *Intensive Care Medicine*. 2020;46(4):586–90.
23. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature Communications*. 2019;10(1).

24. Zou X, Chen K, Zou J, Han P, Hao J, Han Z. Single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to 2019-nCoV infection. *Frontiers of Medicine*. 2020;14(2):185–92.